Journal of Complexity 26 (2010) 629–640

Contents lists available at ScienceDirect

Journal of Complexity

journal homepage: www.elsevier.com/locate/jco

The Gelfand widths of ℓ_p-balls for $0 < p \leq 1$

Simon Foucart [a], Alain Pajor [b], Holger Rauhut [c,*], Tino Ullrich [c]

[a] *Université Pierre et Marie Curie, Laboratoire Jacques-Louis Lions, 175, rue du Chevaleret, 75013 Paris, France*

[b] *Université de Paris Est Marne-la-Vallée, Équipe d'Analyse et de Mathématiques Appliquées 5, Boulevard Descartes, Champs sur Marne, 77454 Marne-la-Vallée Cedex 2, France*

[c] *Hausdorff Center for Mathematics & Institute for Numerical Simulation, University of Bonn, Endenicher Allee 60, 53115 Bonn, Germany*

A R T I C L E I N F O

Article history:
Received 2 February 2010
Accepted 22 April 2010
Available online 10 May 2010

Keywords:
Gelfand widths
Compressive sensing
Sparse recovery
ℓ_1-minimization
ℓ_p-minimization

A B S T R A C T

We provide sharp lower and upper bounds for the Gelfand widths of ℓ_p-balls in the N-dimensional ℓ_q^N-space for $0 < p \leq 1$ and $p < q \leq 2$. Such estimates are highly relevant to the novel theory of compressive sensing, and our proofs rely on methods from this area.

1. Introduction

Gelfand widths are an important concept in classical and modern approximation and complexity theory. They have found recent interest in the rapidly emerging field of compressive sensing [6,14,17] because they give general performance bounds for sparse recovery methods. Since vectors in ℓ_p-balls, $0 < p \leq 1$, can be well-approximated by sparse vectors, the Gelfand widths of such balls are particularly relevant in this context. In remarkable papers [26,21,19] from the 1970s and 1980s due to Kashin, Gluskin, and Garnaev, upper and lower estimates for the Gelfand widths of ℓ_1-balls are provided. In his seminal paper introducing compressive sensing [17], Donoho extends these estimates to the Gelfand widths of ℓ_p-balls with $p < 1$. Unfortunately, his proof of the lower bound contains a gap. In this article, we address this issue by supplying a complete proof. To this end, we proceed in an entirely different way than Donoho. Indeed, we use compressive sensing methods to establish the

* Corresponding author.
E-mail addresses: foucart@ann.jussieu.fr (S. Foucart), alain.pajor@univ-mlv.fr (A. Pajor), rauhut@hcm.uni-bonn.de (H. Rauhut), tino.ullrich@hcm.uni-bonn.de (T. Ullrich).

0885-064X/$ – see front matter
doi:10.1016/j.jco.2010.04.004

lower bound in a more intuitive way. Our method is new even for the case $p = 1$. For completeness, we also give a proof of the upper bound based again on compressive sensing arguments. These arguments also provide the same sharp asymptotic behavior for the Gelfand widths of weak-ℓ_p-balls.

1.1. The main result

In this paper, we consider the finite-dimensional spaces ℓ_p^N, that is, $\mathbb{R}^N$ endowed with the usual ℓ_p-(quasi-)norm defined, for $x \in \mathbb{R}^N$, by

$$\|x\|_p := \left(\sum_{\ell=1}^N |x_\ell|^p\right)^{1/p}, \quad 0 < p < \infty, \qquad \|x\|_\infty := \max_{\ell=1,\dots,N} |x_\ell|.$$

For $1 \le p \le \infty$, this is a norm, while for $0 < p < 1$, it only satisfies the p-triangle inequality

$$\|x+y\|_p^p \le \|x\|_p^p + \|y\|_p^p, \quad x, y \in \mathbb{R}^N. \tag{1.1}$$

Thus, $\|\cdot\|_p$ defines a quasi-norm with (optimal) quasi-norm constant $C = \max\{1, 2^{1/p-1}\}$. As a reminder, $\|\cdot\|_X$ is called a quasi-norm on $\mathbb{R}^N$ with quasi-norm constant $C \ge 1$ if it obeys the quasi-triangle inequality

$$\|x+y\|_X \le C(\|x\|_X + \|y\|_X), \quad x, y \in \mathbb{R}^N.$$

Other quasi-normed spaces considered in this paper are the spaces weak-ℓ_p^N, that is, $\mathbb{R}^N$ endowed with the $\ell_{p,\infty}$-quasi-norm defined, for $x \in \mathbb{R}^N$, by

$$\|x\|_{p,\infty} := \max_{\ell=1,\dots,N} \ell^{1/p}|x_\ell^*|, \quad 0 < p \le \infty,$$

where $x^* \in \mathbb{R}^N$ is a non-increasing rearrangement of x. We shall investigate the Gelfand widths in ℓ_q^N of the unit balls $B_p^N := \{x \in \mathbb{R}^N, \|x\|_p \le 1\}$ and $B_{p,\infty}^N := \{x \in \mathbb{R}^N, \|x\|_{p,\infty} \le 1\}$ of ℓ_p^N and $\ell_{p,\infty}^N$ for $0 < p \le 1$ and $p < q \le 2$.

We recall that the Gelfand width of order m of a subset K of $\mathbb{R}^N$ in the (quasi-)normed space $(\mathbb{R}^N, \|\cdot\|_X)$ is defined as

$$d^m(K, X) := \inf_{A \in \mathbb{R}^{m\times N}} \sup_{v \in K \cap \ker A} \|v\|_X,$$

where $\ker A := \{v \in \mathbb{R}^N, Av = 0\}$ denotes the kernel of A. It is well-known that the above infimum is actually realized [35]. Let us observe that $d^m(K, X) = 0$ for $m \ge N$ when $0 \in K$, so we restrict our considerations to the case $m < N$ in the sequel. Let us also observe that the simple inclusion $B_p^N \subset B_{p,\infty}^N$ implies

$$d^m(B_p^N, \ell_q^N) \le d^m(B_{p,\infty}^N, \ell_q^N).$$

From this point on, we aim at finding a lower bound for $d^m(B_p^N, \ell_q^N)$ and an upper bound for $d^m(B_{p,\infty}^N, \ell_q^N)$ with the same asymptotic behaviors. Our main result is summarized below.

Theorem 1.1. *For $0 < p \le 1$ and $p < q \le 2$, there exist constants $c_{p,q}, C_{p,q} > 0$ depending only on p and q such that, if $m < N$, then*

$$c_{p,q} \min\left\{1, \frac{\ln(N/m)+1}{m}\right\}^{1/p-1/q} \le d^m(B_p^N, \ell_q^N) \le C_{p,q} \min\left\{1, \frac{\ln(N/m)+1}{m}\right\}^{1/p-1/q}, \tag{1.2}$$

and, if $p < 1$,

$$c_{p,q} \min\left\{1, \frac{\ln(N/m)+1}{m}\right\}^{1/p-1/q} \le d^m(B_{p,\infty}^N, \ell_q^N) \le C_{p,q} \min\left\{1, \frac{\ln(N/m)+1}{m}\right\}^{1/p-1/q}. \tag{1.3}$$

In the case $p = 1$ and $q = 2$, the upper bound of (1.2) with a slightly worse log-term was shown by Kashin in [26] by considering Kolmogorov widths, which are dual to Gelfand widths [29,35]. The lower bound and the optimal log-term for $p = 1$ and $1 < q \le 2$ were provided by Garnaev and Gluskin in [21,19], again via Kolmogorov widths. An alternative proof of the upper and lower estimates of (1.2) with $p = 1$ was given by Carl and Pajor in [10]. They did not pass to Kolmogorov widths, but rather used Carl's theorem [9] (see also [11,35]) that bounds in particular Gelfand numbers from below by entropy numbers, which are completely understood even for $p, q < 1$; see [38,25,27]. An upper bound for $p < 1$ and $q = 2$ was first provided by Donoho [17] with $\log(N)$ instead of $\log(N/m)$. With an adaptation of a method from [29], Vybíral [39, Lem. 4.11] also provided the upper bound of (1.2) for $0 < p \le 1$. In Section 3, we use compressive sensing techniques to give an alternative proof that provides the upper bound of (1.3).

Donoho's attempt to prove the lower bound of (1.2) for the case $0 < p < 1$ and $q = 2$ consists in applying Carl's theorem and then using known estimates for entropy numbers, similarly to the approach by Carl and Pajor for $p = 1$. However, it is unknown whether Carl's theorem extends to quasi-norm balls, in particular to ℓ_p-balls with $p < 1$. The standard proof of Carl's theorem for Gelfand widths [11,29] uses duality arguments, which are not available for quasi-Banach spaces. We believe that Carl's theorem actually fails for Gelfand widths of general quasi-norm balls, although it turns out to be a posteriori true in our specific situation due to Theorem 1.1.

We briefly comment on the case $q > 2$. Since then $\|v\|_q \le \|v\|_2$ for all $v \in \mathbb{R}^N$, we have the upper estimate

$$d^m(B_p^N, \ell_q^N) \le d^m(B_p^N, \ell_2^N) \le C_{p,2} \min\left\{1, \frac{\ln(N/m)+1}{m}\right\}^{1/p-1/2}. \tag{1.4}$$

The lower bound in (1.2) extends to $q > 2$, but is unlikely to be optimal in this case. It seems rather that (1.4) is close to the correct behavior. At least for $p = 1$ and $q > 2$, [20] gives lower estimates of related Kolmogorov widths which then lead to (see also [39])

$$d^m(B_1^N, \ell_q^N) \ge c_q m^{-1/2}.$$

The latter matches (1.4) up to the log-factor. We expect a similar behavior for $p < 1$, but this fact remains to be proven.

1.2. The relation to compressive sensing

Let us now outline the connection to compressive sensing. This emerging theory explores the recovery of vectors $x \in \mathbb{R}^N$ from incomplete linear information $y = Ax \in \mathbb{R}^m$, where $A \in \mathbb{R}^{m\times N}$ and $m < N$. Without additional information, reconstruction of x from y is clearly impossible since, even in the full rank case, the system $y = Ax$ has infinitely many solutions. Compressive sensing makes the additional assumption that x is sparse or at least compressible. A vector $x \in \mathbb{R}^N$ is called s-sparse if at most s of its coordinates are non-zero. The error of best s-term approximation is defined as

$$\sigma_s(x)_p := \inf\{\|x - z\|_p, z \text{ is } s\text{-sparse}\}.$$

Informally, a vector x is called compressible if $\sigma_s(x)_p$ decays quickly in s. It is classical to show that, for $q > p$,

$$\sigma_s(x)_q \le \frac{1}{s^{1/p-1/q}} \|x\|_p, \tag{1.5}$$

$$\sigma_s(x)_q \le \frac{D_{p,q}}{s^{1/p-1/q}} \|x\|_{p,\infty}, \qquad D_{p,q} := (q/p-1)^{-1/q}. \tag{1.6}$$

This implies that the balls B_p^N and $B_{p,\infty}^N$ with $p \le 1$ serve as good models for compressible signals: the smaller p, the more closely x in B_p^N or in $B_{p,\infty}^N$ is approximable in ℓ_q^N by s-sparse vectors.

The aim of compressive sensing is to find good pairs of linear measurement maps $A \in \mathbb{R}^{m\times N}$ and (non-linear) reconstruction maps $\Delta : \mathbb{R}^m \to \mathbb{R}^N$ that recover compressible vectors x with small errors

$x - \Delta(Ax)$. In order to measure the performance of a pair (A, Δ), one defines, for a subset K of $\mathbb{R}^N$ and a (quasi-)norm $\|\cdot\|_X$ on $\mathbb{R}^N$,

$$E_m(K, X) := \inf_{A \in \mathbb{R}^{m \times N}, \Delta : \mathbb{R}^m \to \mathbb{R}^N} \sup_{x \in K} \|x - \Delta(Ax)\|_X .$$

Quantities of this type play a crucial role in the modern field of information based complexity [34]. In our situation, only linear information is allowed in order to recover K uniformly. The quantities $E_m(K, X)$ are closely linked to the Gelfand widths, as stated in the following proposition [17,14]; see also [36,33].

Proposition 1.2. *Let $K \subset \mathbb{R}^N$ be such that $K = -K$ and $K + K \subset C_1 K$ for some $C_1 \geq 2$, and let $\|\cdot\|_X$ be a quasi-norm on $\mathbb{R}^N$ with quasi-norm constant C_2. Note that $C_1 = 2$ if K is a norm ball and that $C_2 = 1$ if $\|\cdot\|_X$ is a norm. Then*

$$C_2^{-1} d^m(K, X) \leq E_m(K, X) \leq C_1 d^m(K, X).$$

Combining the previous proposition with Theorem 1.1 gives optimal performance bounds for the recovery of compressible vectors in B_p^N, $0 < p \leq 1$, when the error is measured in ℓ_q, $p < q \leq 2$. Typically, the most interesting case is $q = 2$, for which we end up with

$$c_p \min\left\{1, \frac{\ln(N/m) + 1}{m}\right\}^{1/p - 1/2} \leq E_m(B_p^N, \ell_2^N) \leq C_p \min\left\{1, \frac{\ln(N/m) + 1}{m}\right\}^{1/p - 1/2} .$$

For practical purposes, it is of course desirable to find matrices $A \in \mathbb{R}^{m \times N}$ and efficiently implementable reconstruction maps Δ that realize the optimal bound above. For instance, Gaussian random matrices $A \in \mathbb{R}^{m \times N}$, i.e., matrices whose entries are independent copies of a zero-mean Gaussian variable, provide optimal measurement maps with high probability [8,17,1]. An optimal reconstruction map is obtained via basis pursuit [13,17,8], i.e., via the ℓ_1-minimization mapping given by

$$\Delta_1(y) := \arg\min \|z\|_1 \quad \text{subject to } Az = y.$$

This mapping can be computed with efficient convex optimization methods [2], and works very well in practice. The proof of the lower bound in (1.2) will further involve ℓ_p-minimization for $0 < p \leq 1$ via the mapping

$$\Delta_p(y) := \arg\min \|z\|_p \quad \text{subject to } Az = y.$$

A key concept in the analysis of sparse recovery via ℓ_p-minimization is the restricted isometry property (RIP). This well-established concept in compressive sensing [8,7] is the main tool for the proof of the upper bound in (1.3). We recall that the sth-order restricted isometry constant $\delta_s(A)$ of a matrix $A \in \mathbb{R}^{m \times N}$ is defined as the smallest $\delta > 0$ such that

$$(1 - \delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta)\|x\|_2^2 \quad \text{for all } s\text{-sparse } x \in \mathbb{R}^N.$$

Small restricted isometry constants imply stable recovery by ℓ_1-minimization, as well as by ℓ_p-minimization for $0 < p < 1$. For later reference, we state the following result [7,5,18].

Theorem 1.3. *Suppose $0 < p \leq 1$. If $A \in \mathbb{R}^{m \times N}$ has a restricted isometry constant $\delta_{2s} < \sqrt{2} - 1$, then, for all $x \in \mathbb{R}^N$,*

$$\|x - \Delta_p(Ax)\|_p^p \leq C \sigma_s(x)_p^p, \tag{1.7}$$

where $C > 0$ is a constant that depends only on δ_{2s}. In particular, the reconstruction of s-sparse vectors is exact.

Given a prescribed $0 < \delta < 1$, it is known [1,8,31] that, if the entries of the matrix A are independent copies of a zero-mean Gaussian variable with variance $1/m$, then there exist constants $C_1, C_2 > 0$ (depending only on δ) such that $\delta_s(A) \leq \delta$ holds with probability greater than $1 - e^{-C_2 m}$ provided that

$$m \geq C_1 s \ln(eN/s). \tag{1.8}$$

In particular, there exists a matrix $A \in \mathbb{R}^{m\times N}$ such that the pair (A, Δ_1), and more generally (A, Δ_p) for $0 < p \le 1$, allows stable recovery in the sense of (1.7) as soon as the number of measurements satisfies (1.8). Vice versa, we will see in Theorem 2.7 that the existence of any pair (A, Δ) allowing such a stable recovery forces the number of measurements to satisfy (1.8).

Lemma 2.4, which is of independent interest, estimates the minimal number of measurements for the pair (A, Δ_p) to allow exact (but not necessarily stable) recovery of sparse vectors. Namely, we must have

$$m \ge c_1 ps \ln(N/(c_2 s)) \tag{1.9}$$

for some explicitly given constants $c_1, c_2 > 0$. In the case $p = 1$, this result can be also obtained as a consequence of a corresponding lower bound on neighborliness of centrosymmetric polytopes; see [16,28]. Decreasing p while keeping N fixed shows that the bound (1.9) becomes in fact irrelevant for small p, since the bound $m \ge 2s$ holds as soon as there exists a pair (A, Δ) allowing exact recovery of all s-sparse vectors; see [14, Lem. 3.1]. Combining the two bounds, we see that s-sparse recovery by ℓ_p-minimization forces

$$m \ge C_1 s\,(1 + p\ln(N/(C_2 s)))\,,$$

for some constants $C_1, C_2 > 0$. Interestingly, if such an inequality is fulfilled (with possibly different constants C_1, C_2) and if A is a Gaussian random matrix, then the pair (A, Δ_p) allows s-sparse recovery with high probability; see [12]. We note, however, that exact ℓ_p-minimization with $p < 1$, as a non-convex optimization program, encounters significant difficulties of implementation. For more information on compressive sensing, we refer the reader to [4,6,8,14,17,37].

2. Lower bounds

In this section, we use compressive sensing methods to establish the lower bound in (1.2), and hence the lower bound in (1.3) as a by-product. To be precise, we show the following result, in which the restriction $q \le 2$ is not imposed.

Proposition 2.1. *For $0 < p \le 1$ and $p < q \le \infty$, there exists a constant $c_{p,q} > 0$ such that*

$$d^m(B_p^N, \ell_q^N) \ge c_{p,q} \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1/p-1/q}, \quad m < N. \tag{2.1}$$

The proof of Proposition 2.1 involves several auxiliary steps. We start with a result [23,24] on the unique recovery of sparse vectors via ℓ_p-minimization for $0 < p \le 1$. A proof is included for the reader's convenience. We point out that, given a subset S of $[N] := \{1, \dots, N\}$ and a vector $v \in \mathbb{R}^N$, we denote by v_S the vector that coincides with v on S and that vanishes on the complementary set $S^c := [N] \setminus S$.

Lemma 2.2. *Suppose $0 < p \le 1$ and $N, m, s \in \mathbb{N}$ with $m, s < N$. For a matrix $A \in \mathbb{R}^{m\times N}$, the following statements are equivalent.*

(a) *Every s-sparse vector x is the unique minimizer of $\|z\|_p$ subject to $Az = Ax$.*
(b) *A satisfies the p-null space property of order s, i.e., for every $v \in \ker A \setminus \{0\}$ and every $S \subset [N]$ with $|S| \le s$,*

$$\|v_S\|_p^p < \frac{1}{2}\|v\|_p^p.$$

Proof. (a) $\Rightarrow$ (b): Suppose $v \in \ker A \setminus \{0\}$ and $S \subset [N]$ with $|S| \le s$. Since $v = v_S + v_{S^c}$ satisfies $Av = 0$, we have $Av_S = A(-v_{S^c})$. Then, since v_S is s-sparse, (a) implies

$$\|v_S\|_p^p < \|-v_{S^c}\|_p^p = \|v_{S^c}\|_p^p.$$

Adding $\|v_S\|_p^p$ on both sides and using $\|v_{S^c}\|_p^p + \|v_S\|_p^p = \|v\|_p^p$ gives (b).

(b) $\Rightarrow$ (a): Let x be an s-sparse vector and let $S := \operatorname{supp} x$. Let further $z \neq x$ be such that $Az = Ax$. Then $v := x - z \in \ker A \setminus \{0\}$ and

$$\|x\|_p^p \leq \|x_S - z_S\|_p^p + \|z_S\|_p^p = \|v_S\|_p^p + \|z_S\|_p^p, \tag{2.2}$$

where the first estimate is a consequence of the p-triangle inequality (1.1). Clearly, (b) implies $\|v_S\|_p^p < \|v_{S^c}\|_p^p$. Plugging this into (2.2) and using that $v_{S^c} = -z_{S^c}$ gives

$$\|x\|_p^p < \|v_{S^c}\|_p^p + \|z_S\|_p^p = \|z_{S^c}\|_p^p + \|z_S\|_p^p = \|z\|_p^p.$$

This proves that x is the unique minimizer of $\|z\|_p$ subject to $Az = Ax$. □

The next auxiliary step is a well-known combinatorial lemma; see for instance [32,3,22], [30, Lem. 3.6]. A proof that provides explicit constants is again included for the reader's convenience.

Lemma 2.3. *Suppose $N, s \in \mathbb{N}$ with $s < N$. There exists a family $\mathcal{U}$ of subsets of $[N]$ such that:*

(i) *Every set in $\mathcal{U}$ consists of exactly s elements.*
(ii) *For all $I, J \in \mathcal{U}$ with $I \neq J$, it holds that $|I \cap J| < s/2$.*
(iii) *The family $\mathcal{U}$ is "large" in the sense that*

$$|\mathcal{U}| \geq \left(\frac{N}{4s}\right)^{s/2}.$$

Proof. We may assume that $s \leq N/4$, for otherwise we can take a family $\mathcal{U}$ consisting of just one element. Let us denote by $\mathcal{B}(N, s)$ the family of subsets of $[N]$ having exactly s elements. This family has size $|\mathcal{B}(N, s)| = \binom{N}{s}$. We draw an arbitrary element $I_1 \in \mathcal{B}(N, s)$ and collect in a family $\mathcal{A}_1$ all the sets $J \in \mathcal{B}(N, s)$ such that $|I \cap J| \geq s/2$. Then $\mathcal{A}_1$ has size at most

$$\sum_{k=\lceil s/2\rceil}^{s} \binom{s}{k}\binom{N-s}{s-k} \leq 2^s \max_{\lceil s/2\rceil \leq k \leq s} \binom{N-s}{s-k} = 2^s \binom{N-s}{\lfloor s/2\rfloor}, \tag{2.3}$$

the latter inequality holding because $\lfloor s/2\rfloor \leq (N-s)/2$ when $s \leq N/2$. We throw away $\mathcal{A}_1$ and observe that every element in $J \in \mathcal{B}(N, s) \setminus \mathcal{A}_1$ satisfies $|I_1 \cap J| < s/2$. Next we draw an arbitrary element $I_2 \in \mathcal{B}(N, s) \setminus \mathcal{A}_1$, provided that the latter is not empty. We repeat the procedure, i.e., we define a family $\mathcal{A}_2$ relative to I_2 and draw an arbitrary element $I_3 \in \mathcal{B}(N, s) \setminus (\mathcal{A}_1 \cup \mathcal{A}_2)$, and so forth until no more elements are left. The size of each set $\mathcal{A}_i$ can always be estimated from above by (2.3). This results in a collection $\mathcal{U} = \{I_1, \ldots, I_L\}$ of subsets of $[N]$ satisfying (i) and (ii). We finally observe that

$$\begin{aligned} L &\geq \frac{\binom{N}{s}}{2^s \binom{N-s}{\lfloor s/2\rfloor}} = \frac{1}{2^s} \frac{N(N-1)\cdots(N-s+1)}{(N-s)(N-s-1)\cdots(N-s-\lfloor s/2\rfloor+1)} \frac{1}{s(s-1)\cdots(\lfloor s/2\rfloor+1)} \\ &\geq \frac{1}{2^s} \frac{N(N-1)\cdots(N-\lceil s/2\rceil+1)}{s(s-1)\cdots(s-\lceil s/2\rceil+1)} \geq \frac{1}{2^s}\left(\frac{N}{s}\right)^{\lceil s/2\rceil} \geq \left(\frac{N}{4s}\right)^{s/2}. \end{aligned}$$

This concludes the proof by establishing (iii). □

We now use Lemma 2.3 for the final auxiliary result, which is quite interesting on its own. It gives an estimate of the minimal number of measurements for exact recovery of sparse vectors via ℓ_p-minimization, where $0 < p \leq 1$.

Lemma 2.4. *Suppose $0 < p \leq 1$ and $N, m, s \in \mathbb{N}$ with $m < N$ and $s < N/2$. If $A \in \mathbb{R}^{m\times N}$ is a matrix such that every $2s$-sparse vector x is a minimizer of $\|z\|_p$ subject to $Az = Ax$, then*

$$m \geq c_1 p s \ln\left(\frac{N}{c_2 s}\right),$$

where $c_1 := 1/\ln 9 \approx 0.455$ and $c_2 := 4$.

Remark 2.5. Lemma 2.4 could be rephrased (with modified constants) by replacing $2s$-sparse vectors, $s \geq 1$, by s-sparse vectors, $s \geq 2$. In the case $s = 1$, it is possible for every 1-sparse vector x to be a (nonunique) minimizer of $\|z\|_1$ subject to $Az = Ax$, yet $m \geq c_1 p \ln(N/c_2)$ fails for all constants $c_1, c_2 > 0$. This can be verified by taking $m = 1$ and $A = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$.

Proof. We consider the quotient space

$$X := \mathbb{R}^N / \ker A = \{[x] := x + \ker A,\ x \in \mathbb{R}^N\},$$

which has algebraic dimension $r := \mathrm{rank} A \leq m$. It is a quasi-Banach space equipped with

$$\|[x]\|_{A,p} := \inf_{v \in \ker A} \|x + v\|_p.$$

Indeed, a simple computation reveals that $\|\cdot\|_{A,p}$ satisfies the p-triangle inequality, i.e.,

$$\|[x] + [y]\|_{A,p}^p \leq \|[x]\|_{A,p}^p + \|[y]\|_{A,p}^p.$$

By assumption, the quotient map $[\cdot]$ preserves the norm of every $2s$-sparse vector. We now choose a family $\mathcal{U}$ of subsets of $[N]$ satisfying (i)–(iii) of Lemma 2.3. For a set $I \in \mathcal{U}$, we define an element $x_I \in \ell_p^N$ with $\|x_I\|_p = 1$ by

$$x_I := \frac{1}{s^{1/p}} \sum_{i \in I} e_i, \tag{2.4}$$

where $(e_1, \dots, e_N)$ denotes the canonical basis of $\mathbb{R}^N$. For $I, J \in \mathcal{U}$, $I \neq J$, (ii) yields

$$\|x_I - x_J\|_p^p > \frac{2s - 2s/2}{s} = 1.$$

Since the vector $x_I - x_J$ is a $2s$-sparse vector, we obtain

$$\|[x_I] - [x_J]\|_{A,p} = \|[x_I - x_J]\|_{A,p} = \|x_I - x_J\|_p > 1.$$

The p-triangle inequality implies that $\{[x_I] + (1/2)^{1/p} B_X, I \in \mathcal{U}\}$ is a disjoint collection of balls included in the ball $(3/2)^{1/p} B_X$, where B_X denotes the unit ball of $(X, \|\cdot\|_{A,p})$. Let $\mathrm{vol}(\cdot)$ denote a volume form on X, that is a translation invariant measure satisfying $\mathrm{vol}(B_X) > 0$ and $\mathrm{vol}(\lambda B_X) = \lambda^r \mathrm{vol}(\lambda B_X)$ for all $\lambda > 0$ (such a measure exists since X is isomorphic to $\mathbb{R}^r$). The volumes satisfy the relation

$$\sum_{I \in \mathcal{U}} \mathrm{vol}\left([x_I] + (1/2)^{1/p} B_X\right) \leq \mathrm{vol}\left((3/2)^{1/p} B_X\right).$$

By translation invariance and homogeneity, we then derive

$$|\mathcal{U}| (1/2)^{r/p} \mathrm{vol}(B_X) \leq (3/2)^{r/p} \mathrm{vol}(B_X).$$

As a result of (iii), we finally obtain

$$\left(\frac{N}{4s}\right)^{s/2} \leq 3^{r/p} \leq 3^{m/p}.$$

Taking the logarithm on both sides gives the desired result. □

Now we are ready to prove Proposition 2.1. The underlying idea is that a small Gelfand width would imply $2s$-sparse recovery for s large enough to violate the conclusion of Lemma 2.4.

Proof. With $c := (1/2)^{2/p - 1/q}$ and $d := 2c_1 p/(4 + c_1) \approx 0.204p$, we are going to prove that

$$d^m(B_p^N, \ell_q^N) \geq c \mu^{1/p - 1/q}, \quad \text{where } \mu := \min\left\{1, \frac{d \ln(eN/m)}{m}\right\}. \tag{2.5}$$

The desired result will follow with $c_{p,q} := cd^{1/p-1/q}$. By way of contradiction, we assume that $d^m(B_p^N, \ell_q^N) < c\mu^{1/p-1/q}$. This implies the existence of a matrix $A \in \mathbb{R}^{m\times N}$ such that, for all $v \in \ker A \setminus \{0\}$,

$$\|v\|_q < c\mu^{1/p-1/q}\|v\|_p.$$

For a fixed $v \in \ker A \setminus \{0\}$, in view of the inequalities $\|v\|_p \le N^{1/p-1/q}\|v\|_q$ and $c \le (1/2)^{1/p-1/q}$, we derive $1 < (\mu N/2)^{1/p-1/q}$, so $1 \le 1/\mu < N/2$. We then define $s := \lfloor 1/\mu \rfloor \ge 1$, so

$$\frac{1}{2\mu} < s \le \frac{1}{\mu}.$$

Now, for $v \in \ker A \setminus \{0\}$ and $S \subset [N]$ with $|S| \le 2s$, we have

$$\|v_S\|_p \le (2s)^{1/p-1/q}\|v_S\|_q \le (2s)^{1/p-1/q}\|v\|_q < c(2s\mu)^{1/p-1/q}\|v\|_p \le \frac{1}{2^{1/p}}\|v\|_p.$$

This shows that the p-null space property of order $2s$ is satisfied. Hence, Lemmas 2.2 and 2.4 imply

$$m \ge c_1 ps \ln\left(\frac{N}{c_2 s}\right). \tag{2.6}$$

Besides, since the pair (A, Δ_p) allows exact recovery of all $2s$-sparse vectors, we have

$$m \ge 2(2s) = c_2 s. \tag{2.7}$$

Using (2.7) in (2.6), it follows that

$$m \ge c_1 ps \ln\left(\frac{N}{m}\right) = c_1 ps \ln\left(\frac{eN}{m}\right) - c_1 ps > \frac{c_1 p}{2\mu}\ln\left(\frac{eN}{m}\right) - \frac{c_1}{4}m.$$

After rearrangement, we deduce

$$m > \frac{2c_1 p}{4+c_1}\frac{\ln(eN/m)}{\min\{1, d\ln(eN/m)/m\}} \ge \frac{2c_1 p}{4+c_1}\frac{\ln(eN/m)}{d\ln(eN/m)}m = m.$$

This is the desired contradiction. $\square$

Remark 2.6. When m is close to N, the lower estimate (2.5) is rather poor. In this case, a nice and simple argument proposed to us by Vybíral gives the improved estimate

$$d^m(B_p^N, \ell_q^N) \ge \left(\frac{1}{m+1}\right)^{1/p-1/q}, \quad m < N. \tag{2.8}$$

Indeed, for an arbitrary matrix $A \in \mathbb{R}^{m\times N}$, the kernel of A and the $(m+1)$-dimensional space $\{x \in \mathbb{R}^N : x_i = 0 \text{ for all } i > m+1\}$ have a nontrivial intersection. We then choose a vector $v \ne 0$ in this intersection, and (2.8) follows from the inequality $\|v\|_p \le (m+1)^{1/p-1/q}\|v\|_q$.

We close this section with the important observation that any measurement/reconstruction scheme that provides ℓ_1-stability requires a number of measurements scaling at least like the sparsity times a log-term. This may be viewed as a consequence of Propositions 1.2 and 2.1. Indeed, fixing $p < 1$, the inequalities (1.5) and (1.7) imply

$$d^m(B_p^N, \ell_1^N) \le E_m(B_p^N, \ell_1^N) \le C \sup_{x\in B_p^N} \sigma_s(x)_1 \le \frac{C}{s^{1/p-1}}.$$

The lower bound (2.1) for the Gelfand width then yields, for some constant c,

$$c\min\left\{1, \frac{\ln(eN/m)}{m}\right\} \le \frac{1}{s}.$$

We derive either $s \leq 1/c$ or $m \geq cs\ln(eN/m)$. In short, if $s > 1/c$, then ℓ_1-stability implies $m \geq cs\ln(eN/m)$ — which can be shown to imply in turn $m \geq c's\ln(eN/s)$. We provide below a direct argument that removes the restriction $s > 1/c$. It uses Lemma 2.3 and works also for ℓ_p-stability with $p < 1$. It borrows ideas from a paper by Do Ba et al. [15, Thm. 3.1], which contains the case $p = 1$ in a stronger non-uniform version.

Theorem 2.7. *Suppose $N, m, s \in \mathbb{N}$ with $m, s < N$. Suppose that a measurement matrix $A \in \mathbb{R}^{m\times N}$ and a reconstruction map $\Delta : \mathbb{R}^N \to \mathbb{R}^m$ are stable in the sense that, for all $x \in \mathbb{R}^N$,*

$$\|x - \Delta(Ax)\|_p^p \leq C\sigma_s(x)_p^p$$

for some constant $C > 0$ and some $0 < p \leq 1$. Then there exists a constant $C' > 0$ depending only on C such that

$$m \geq C'ps\ln(eN/s).$$

Proof. We consider again a family $\mathcal{U}$ of subsets of $[N]$ given by Lemma 2.3. For each $I \in \mathcal{U}$, we define an s-sparse vector x_I with $\|x_I\|_p = 1$ as in (2.4). With $\rho := (2(C+1))^{-1/p}$, we claim that $\{A(x_I + \rho B_p^N), I \in \mathcal{U}\}$ is a disjoint collection of subsets of $A(\mathbb{R}^N)$, which has algebraic dimension $r \leq m$. Suppose indeed that there exist $I, J \in \mathcal{U}$ with $I \neq J$ and $z, z' \in \rho B_p^N$ such that $A(x_I + z) = A(x_J + z')$. A contradiction follows from

$$\begin{aligned} 1 < \|x_I - x_J\|_p^p &\leq \|x_I + z - \Delta(A(x_I+z))\|_p^p + \|x_J + z' - \Delta(A(x_J + z'))\|_p^p + \|z\|_p^p + \|z'\|_p^p \\ &\leq C\sigma_s(x_I+z)_p^p + C\sigma_s(x_J+z')_p^p + \|z\|_p^p + \|z'\|_p^p \\ &\leq C\|z\|_p^p + C\|z'\|_p^p + \|z\|_p^p + \|z'\|_p^p \leq 2(C+1)\rho^p = 1. \end{aligned}$$

We now observe that the collection $\{A(x_I + \rho B_p^N), I \in \mathcal{U}\}$ is contained in $(1+\rho^p)^{1/p}A(B_p^N)$. As in the proof of Lemma 2.4, we use a standard volumetric argument to derive

$$|\mathcal{U}|\rho^r \mathrm{vol}(A(B_p^N)) = \sum_{I\in\mathcal{U}} \mathrm{vol}(A(x_I + \rho B_p^N)) \leq \mathrm{vol}((1+\rho^p)^{1/p}A(B_p^N)) = (1+\rho^p)^{r/p}\mathrm{vol}(A(B_p^N)).$$

We deduce that

$$\left(\frac{N}{4s}\right)^{s/2} \leq (\rho^{-p}+1)^{r/p} \leq (\rho^{-p}+1)^{m/p} = (2C+3)^{m/p}.$$

Taking the logarithm on both sides yields

$$m \geq cps\ln(N/(4s)), \quad \text{with } c := 1/(2\ln(2C+3)).$$

Finally, noticing that $m \geq 2s$ because the pair (A, Δ) allows exact s-sparse recovery, we obtain

$$m \geq cps\ln(eN/s) - cps\ln(4e) \geq cps\ln(eN/s) - \frac{c\ln(4e)}{2}m.$$

The desired result follows with $C' := (2c)/(2 + c\ln(4e))$. $\square$

3. Upper bounds

In this section, we establish the upper bound in (1.3), and hence the upper bound in (1.2) as a by-product. As already mentioned in the introduction, the bound for the Gelfand width of ℓ_p-balls was already provided by Vybíral in [39], but the bound for the Gelfand width of weak-ℓ_p-balls is new.

Proposition 3.1. *For $0 < p < 1$ and $p < q \leq 2$, there exists a constant $C_{p,q} > 0$ such that*

$$d^m(B_{p,\infty}^N, \ell_q^N) \leq C_{p,q}\min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1/p-1/q}, \quad m < N. \tag{3.1}$$

The argument relies again on compressive sensing methods. According to Proposition 1.2, it is enough to establish the upper bound for the quantity $E_m(B^N_{p,\infty}, \ell^N_q)$. This is done in the following theorem, which we find rather illustrative because it shows that, even when $p < 1$, an optimal reconstruction map Δ for the realization of the number $E_m(B^N_{p,\infty}, \ell^N_q)$ can be chosen to be the ℓ_1-minimization mapping, at least when $q \geq 1$. The argument is originally due to Donoho for $q = 2$ [17, Proof of Theorem 9] and can be extended to all $2 \geq q > p$.

Theorem 3.2. *For $0 < p < 1$ and $p < q \leq 2$, there exists a matrix $A \in \mathbb{R}^{m\times N}$ such that, with $r = \min\{1, q\}$,*

$$\sup_{x\in B^N_{p,\infty}} \|x - \Delta_r(Ax)\|_q \leq C_{p,q} \min\left\{1, \frac{\ln(N/m)+1}{m}\right\}^{1/p-1/q},$$

where $C_{p,q} > 0$ is a constant that depends only on p and q.

Proof. Let C_1 be the constant in (1.8) relative to the RIP associated with $\delta = 1/3$, say. We choose a constant $D > 0$ large enough to have

$$D/2 > e, \quad \frac{D/2}{1+\ln(D/2)} > C_1.$$

We are going to prove that, for any $x \in B^N_{p,\infty}$,

$$\|x - \Delta_r(Ax)\|_q \leq C'_{p,q} \min\left\{1, \frac{D\ln(eN/m)}{m}\right\}^{1/p-1/q} \tag{3.2}$$

for some constant $C'_{p,q} > 0$. This will imply the desired result with $C_{p,q} := C'_{p,q} D^{1/p-1/q}$.

Case 1: $m > D\ln(eN/m)$.

We define $s \geq 1$ as the largest integer smaller than $m/(D\ln(eN/m))$, so

$$\frac{m}{2D\ln(eN/m)} \leq s < \frac{m}{D\ln(eN/m)}. \tag{3.3}$$

Putting $t = 2s$ and noticing that $t/m < 2/D < 1/e$ and that $u \mapsto u\ln(u)$ is decreasing on $[0, 1/e]$, we obtain

$$m > \frac{D}{2} t\ln(eN/m) = \frac{D}{2} t\ln(eN/t) + \frac{D}{2} m(t/m)\ln(t/m) > \frac{D}{2} t\ln(eN/t) - m\ln(D/2),$$

and so

$$m > \frac{D/2}{1+\ln(D/2)} t\ln(eN/t) > C_1 t\ln(eN/t).$$

It is then possible to find a matrix $A \in \mathbb{R}^{m\times N}$ with $\delta_t(A) \leq \delta$. In particular, we have $\delta_s(A) \leq \delta$. Now, given $v := x - \Delta_r(Ax) \in \ker A$, we decompose $[N]$ as the disjoint union of sets $S_1, S_2, S_3, \ldots$ of size s (except maybe the last one) in such a way that $|v_i| \geq |v_j|$ for all $i \in S_{k-1}, j \in S_k$, and $k \geq 2$. This easily implies $(\|v_{S_k}\|_2^2/s)^{1/2} \leq (\|v_{S_{k-1}}\|_r^r/s)^{1/r}$, i.e.,

$$\|v_{S_k}\|_2 \leq \frac{1}{s^{1/r-1/2}} \|v_{S_{k-1}}\|_r, \quad k \geq 2. \tag{3.4}$$

Using the r-triangle inequality, we have

$$\|v\|_q^r = \left\|\sum_{k\geq 1} v_{S_k}\right\|_q^r \leq \sum_{k\geq 1} \|v_{S_k}\|_q^r \leq \sum_{k\geq 1} \left(s^{1/q-1/2}\|v_{S_k}\|_2\right)^r \leq \sum_{k\geq 1} \left(\frac{s^{1/q-1/2}}{\sqrt{1-\delta}}\|Av_{S_k}\|_2\right)^r.$$

The fact that $v \in \ker A$ implies $Av_{S_1} = -\sum_{k\geq 2} Av_{S_k}$. It follows that

$$\|v\|_q^r \leq \left(\frac{s^{1/q-1/2}}{\sqrt{1-\delta}}\right)^r \left(\sum_{k\geq 2} \|Av_{S_k}\|_2\right)^r + \left(\frac{s^{1/q-1/2}}{\sqrt{1-\delta}}\right)^r \sum_{k\geq 2} \|Av_{S_k}\|_2^r$$
$$\leq 2\left(\frac{s^{1/q-1/2}}{\sqrt{1-\delta}}\right)^r \sum_{k\geq 2} \|Av_{S_k}\|_2^r \leq 2\left(\sqrt{\frac{1+\delta}{1-\delta}}s^{1/q-1/2}\right)^r \sum_{k\geq 2} \|v_{S_k}\|_2^r.$$

We then derive, using the inequality (3.4),

$$\|v\|_q^r \leq 2\left(\sqrt{\frac{1+\delta}{1-\delta}}\frac{1}{s^{1/r-1/q}}\right)^r \sum_{k\geq 1} \|v_{S_k}\|_r^r.$$

In view of the choice $\delta = 1/3$ and of (3.3), we deduce

$$\|x - \Delta_r(Ax)\|_q \leq 2^{1/r}\sqrt{2}\left(\frac{2D\ln(eN/m)}{m}\right)^{1/r-1/q} \|x - \Delta_r(Ax)\|_r. \tag{3.5}$$

Moreover, in view of $\delta_{2s} \leq 1/3$ and of Theorem 1.3, there exists a constant $C > 0$ such that

$$\|x - \Delta_r(Ax)\|_r \leq C^{1/r}\sigma_s(x)_r. \tag{3.6}$$

Finally, using (1.6) and (3.3), we have

$$\sigma_s(x)_r \leq \frac{D_{p,r}}{s^{1/p-1/r}} \leq D_{p,r}\left(\frac{2D\ln(eN/m)}{m}\right)^{1/p-1/r}. \tag{3.7}$$

Putting (3.5)–(3.7) together, we obtain, for any $x \in B_{p,\infty}^N$,

$$\|x - \Delta_r(Ax)\|_q \leq C''_{p,q}\left(\frac{D\ln(eN/m)}{m}\right)^{1/p-1/q} = C''_{p,q}\min\left\{1, \frac{D\ln(eN/m)}{m}\right\}^{1/p-1/q},$$

where $C''_{p,q} := C^{1/r}D_{p,r}2^{1/r+1/2+1/p-1/q}$.

Case 2: $m \leq D\ln(eN/m)$.

We simply choose the matrix $A \in \mathbb{R}^{m\times N}$ as $A = 0$. Then, for any $x \in B_{p,\infty}^N$, we have

$$\|x - \Delta_r(Ax)\|_q = \|x\|_q \leq C'''_{p,q}\|x\|_{p,\infty} \leq C'''_{p,q},$$

for some constant $C'''_{p,q} > 0$. This yields

$$\|x - \Delta_r(Ax)\|_q \leq C'''_{p,q}\min\left\{1, \frac{D\ln(eN/m)}{m}\right\}^{1/p-1/q}.$$

Both cases show that (3.2) is valid with $C'_{p,q} := \max\{C''_{p,q}, C'''_{p,q}\}$. This completes the proof. $\square$

Remark 3.3. The case $p = 1$, for which $r = 1$, is not covered by our arguments. Since $\sup_{x\in B_{1,\infty}^N} \sigma_s(x)_1 \asymp \log(N/s)$ the quantity $\sigma_s(x)_1$ cannot be bounded by a constant times $\|x\|_{1,\infty}$ in order to obtain (3.7). Instead, the additional log-factor $\log(N/m)$ appears on the right-hand side and therefore in the upper estimate of (1.3) in the case $p = 1$. The correct behavior of the Gelfand widths of weak-ℓ_1-balls does not seem to be known. Nonetheless, the inequality $\sigma_s(x)_1 \leq \|x\|_1$ is always true. This yields the well-known upper estimate for the Gelfand widths of ℓ_1-balls and hence completes the proof of Theorem 1.1.

Acknowledgments

The first author is supported by the French National Research Agency (ANR) through the project ECHANGE (ANR-08-EMER-006). The third and fourth authors acknowledge support by the Hausdorff Center for Mathematics, University of Bonn. The third author acknowledges funding through the WWTF project SPORTS (MA07-004).

References

[1] R.G. Baraniuk, M. Davenport, R.A. DeVore, M. Wakin, A simple proof of the restricted isometry property for random matrices, Constr. Approx. 28 (3) (2008) 253–263.
[2] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge Univ. Press, 2004.
[3] H. Buhrman, P. Miltersen, J. Radhakrishnan, S. Venkatesh, Are bitvectors optimal? in: STOC'00: Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, ACM, 2000, pp. 449–458.
[4] E.J. Candès, Compressive sampling, in: Proceedings of the International Congress of Mathematicians, 2006.
[5] E.J. Candès, The restricted isometry property and its implications for compressed sensing, C. R. Acad. Sci., Paris Sér. I Math. 346 (2008) 589–592.
[6] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inform. Theory 52 (2) (2006) 489–509.
[7] E.J. Candès, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm. Pure Appl. Math. 59 (8) (2006) 1207–1223.
[8] E.J. Candès, T. Tao, Near optimal signal recovery from random projections: universal encoding strategies? IEEE Trans. Inform. Theory 52 (12) (2006) 5406–5425.
[9] B. Carl, Entropy numbers, s-numbers, and eigenvalue problems, J. Funct. Anal. 41 (1981) 290–306.
[10] B. Carl, A. Pajor, Gel'fand numbers of operators with values in a Hilbert space, Invent. Math. 94 (3) (1988) 479–504.
[11] B. Carl, I. Stephani, Entropy, Compactness and the Approximation of Operators, Cambridge University Press, 1990.
[12] R. Chartrand, V. Staneva, Restricted isometry properties and nonconvex compressive sensing, Inverse Problems 24 (035020) (2008) 1–14.
[13] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM J. Sci. Comput. 20 (1) (1999) 33–61.
[14] A. Cohen, W. Dahmen, R.A. DeVore, Compressed sensing and best k-term approximation, J. Amer. Math. Soc. 22 (1) (2009) 211–231.
[15] K. Do Ba, P. Indyk, E. Price, D. Woodruff, Lower bounds for sparse recovery, in: Proc. SODA, 2010.
[16] D.L. Donoho, Neighborly polytopes and sparse solutions of underdetermined linear equations, 2005, preprint.
[17] D.L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306.
[18] S. Foucart, M. Lai, Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$, Appl. Comput. Harmon. Anal. 26 (3) (2009) 395–407.
[19] A.Y. Garnaev, E.D. Gluskin, On widths of the Euclidean ball, Soviet Math. Dokl. 30 (1984) 200–204.
[20] E.D. Gluskin, On some finite-dimensional problems in the theory of widths, Vestn. Leningr. Univ. Math. 14 (1982) 163–170.
[21] E.D. Gluskin, Norms of random matrices and widths of finite-dimensional sets, Math. USSR-Sb. 48 (1984) 173–182.
[22] R. Graham, N. Sloane, Lower bounds for constant weight codes, IEEE Trans. Inform. Theory 26 (1) (1980) 37–43.
[23] R. Gribonval, M. Nielsen, Sparse representations in unions of bases, IEEE Trans. Inform. Theory 49 (12) (2003) 3320–3325.
[24] R. Gribonval, M. Nielsen, Highly sparse representations from dictionaries are unique and independent of the sparseness measure, Appl. Comput. Harmon. Anal. 22 (3) (2007) 335–355.
[25] O. Guédon, A.E. Litvak, Euclidean projections of a p-convex body, in: Geometric Aspects of Functional Analysis, in: Lecture Notes in Math., vol. 1745, Springer, Berlin, 2000.
[26] B.S. Kashin, Diameters of some finite-dimensional sets and classes of smooth functions, Math. USSR. Izv. 11 (1977) 317–333.
[27] T. Kühn, A lower estimate for entropy numbers, J. Approx. Theory 110 (1) (2001) 120–124.
[28] N. Linial, I. Novik, How neighborly can a centrally symmetric polytope be? Discrete Comput. Geom. 36 (2) (2006) 273–281.
[29] G. Lorentz, M. von Golitschek, Y. Makovoz, Constructive Approximation: Advanced Problems, in: Grundlehren der Mathematischen Wissenschaften, vol. 304, Springer, Berlin, 1996.
[30] S. Mendelson, A. Pajor, M. Rudelson, The geometry of random $\{-1, 1\}$-polytopes, Discrete Comput. Geom. 34 (3) (2005) 365–379.
[31] S. Mendelson, A. Pajor, N. Tomczak-Jaegermann, Uniform uncertainty principle for Bernoulli and subgaussian ensembles, Constr. Approx. 28 (3) (2009) 277–289.
[32] N. Noam, W. Avi, Hardness vs randomness, J. Comput. System Sci. 49 (2) (1994) 149–167.
[33] E. Novak, Optimal recovery and n-widths for convex classes of functions, J. Approx. Theory 80 (3) (1995) 390–408.
[34] E. Novak, H. Woźniakowski, Tractability of Multivariate Problems. Vol. I. Linear Information, EMS, Zürich, 2008.
[35] A. Pinkus, N-Widths in Approximation Theory, Springer-Verlag, Berlin, 1985.
[36] A. Pinkus, N-widths and Optimal Recovery, Lect. Notes AMS Short Course ed., in: Proc. Symp. Appl. Math., vol. 36, 1986, pp. 51–66.
[37] H. Rauhut, Compressive sensing and structured random matrices, in: M. Fornasier (Ed.), Theoretical Foundations and Numerical Methods for Sparse Recovery, Radon Series Comp. Appl. Math. 9, De Gruyter, 2010.
[38] C. Schütt, Entropy numbers of diagonal operators between symmetric Banach spaces, J. Approx. Theory 40 (2) (1984) 121–128.
[39] J. Vybíral, Widths of embeddings in function spaces, J. Complexity 24 (4) (2008) 545–570.